COMPLIANCE COMPONENT

DEFINITION

Name	Page Ranking Criteria
Description	Search engines utilize methods of determining the best possible results for searches. This method is called page ranking. It is the numerical weighting given to hyperlinked documents (or Web pages) indexed by a search engine. All search engines have basic rules they must follow which make them distinct in how they use their proprietary algorithms when creating indices for searches.
Rationale	How you construct a page can influence its page ranking. How you use the basic rules can boost or hinder your Web page rankings.
Benefits	Benefits include: • Users finding what they need quicker; • Users finding satisfactory results within fewer clicks; and • Utilizing basic rules can give Web pages better page ranking results, therefore returning the best possible search results for visitors to a site.

ASSOCIATED ARCHITECTURE LEVELS

Specify the Domain Name	Interface
Specify the Discipline Name	Branding
Specify the Technology Area Name	Search Engines
Specify the Product Component Name	

COMPLIANCE COMPONENT TYPE

Document the Compliance Component Type	Guideline
Component Sub-type	

COMPLIANCE DETAIL

State the Guideline, Standard or Legislation	Listed below are guidelines for making your Web site search friendly. • Order of importance – page title, headings, visible content, Meta tags (description and keywords), ALT attributes and then hidden comments. Note: some search engines ignore the hidden comments for ranking purpose, order of importance and weighting is determined by each search engine and may not be consistent across those available. • Make a site with a clear hierarchy and text hyperlinks. Every public page should be reachable from at least one static text hyperlink. • Offer a site map to your users with hyperlinks that point to the important parts of your site. If the site map is larger than 100 hyperlinks, you may want to break the site map into separate pages. • Think about the words users would use to find your pages, and make sure your site includes those words within it, both in the content and Meta tag (keywords). • Try to use text instead of images to display important names, content, or hyperlinks. The search engine crawler doesn't recognize graphic

representation of text contained in the image; instead it looks at the ALT attribute only.

- Use keywords within ALT attributes since search engines use them to rank.
- Make sure that your TITLE and ALT attributes are descriptive and accurate.
- Check for broken hyperlinks and correct markup language.
- If you decide to use dynamic pages, i.e., the URL (uniform resource locator) contains a "?" character, be aware that not every search engine spider crawls dynamic pages as well as it does static pages. It helps to keep the parameters short and the number of them few.
- Keep the hyperlinks on a given page to a reasonable number (fewer than 100).
- For static content, use the Meta tags to help the search robot index only your permanent, non-changing hyperlinks called permalinks not your constantly changing hyperlinks. To do this, use <meta name="robots" content="noindex,follow"> on your front page and <meta name="robots" content="index,follow"> on your secondary pages. INDEX means to include the page in the index the search engine is creating. NOINDEX tells the search engine to not include the page in the index. It is this index that the search engine uses to find user search results. If the page isn't added to the index then it will not be found. FOLLOW indicates the search engine should follow the hyperlinks on your page, and NOFOLLOW tells the search engine not to follow hyperlinks found on the page.
- Make use of the robots.txt file on your Web server. This file tells crawlers which directories can or cannot be crawled. Make sure it is current for your site so you don't accidentally block the search engine crawler. This file needs to be located in the root level of your domain space. A few examples are listed below. The '*' (wildcard) in the User-agent field is a special value meaning "any robot". You need a separate "Disallow" line for every URL prefix you want to exclude. Blank lines are not allowed, as they are used to delimit multiple records. Everything not explicitly disallowed is considered fair game to retrieve.
  To find out more regarding the use of "robots.txt", visit
  http://www.robotstxt.org/wc/exclusion.html.
- To find out what criteria is used when the search engine's rank Web pages, visit http://www.searchengines.com/search_engine_statistics.html.

Document Source Reference #	

Compliance Sources			
Name	Google Information for Webmasters	*Website*	http://www.google.com/webmasters/guidelines.html
Contact Information			
Name	Yahoo Search Basics	*Website*	http://help.yahoo.com/help/us/ysearch/ranking/ranking-02.html
Contact Information			

KEYWORDS	
List Keywords	Page Ranking, Search Engine, Robot (Bot), Web Crawler, Spider, Meta tags, Googlebot, criteria, results, index, indexing, hyperlinks, algorithms, robots.txt

COMPONENT CLASSIFICATION				
Provide the Classification	☐ *Emerging*	☒ *Current*	☐ *Twilight*	☐ *Sunset*

Sunset Date		

COMPONENT SUB-CLASSIFICATION		
Sub-Classification	Date	Additional Sub-Classification Information
☐ Technology Watch		
☐ Variance		
☐ Conditional Use		

Rationale for Component Classification	
Document the Rationale for Component Classification	

Migration Strategy	
Document the Migration Strategy	

Impact Position Statement	
Document the Position Statement on Impact	

CURRENT STATUS	
Provide the Current Status	☐ In Development ☒ Under Review ☐ Approved ☐ Rejected

AUDIT TRAIL			
Creation Date	2/6/2006	Date Approved / Rejected	7/11/06
Reason for Rejection			
Last Date Reviewed		Last Date Updated	
Reason for Update			